MAX PLANCK INSTITUTE
FOR DYNAMICS OF COMPLEX
TECHNICAL SYSTEMS
MAGDEBURG

CSC
COMPUTATIONAL METHODS IN
SYSTEMS AND CONTROL THEORY

Tensor-space Galerkin POD for (optimal control of) parametric flow equations

Peter Benner **Jan Heiland**

July 18

Valencia, ICIAM 2019
Profs. Quaini&Chen&Rozza's MS:
ROM for Parametric CFD Problems

Outline

$\dot{x} - \Delta x = f$

Consider the solution of a PDE:

$$x \in L^2(I; L^2(\Omega))$$

with $I \subset \mathbb{R}$ ... the time-interval
$\Omega \subset \mathbb{R}^n$ ... the spatial domain

and its numerical approximation:

$$\mathbf{x} \in \mathcal{S} \cdot \mathcal{Y}$$

with $\mathcal{S} \subset L^2(I)$ ... discretized time
$\mathcal{Y} \subset L^2(\Omega)$ ... a FE space

Task: Find $\hat{\mathcal{S}} \subset \mathcal{S}$ and $\hat{\mathcal{Y}} \subset \mathcal{Y}$ of much smaller dimension to express $\mathbf{x}$.

Space-Time Spaces

PDE solution $x \in L^2(I; L^2(\Omega))$
$\mathcal{S} \subset L^2(I)$ ... discretized time
$\mathcal{Y} \subset L^2(\Omega)$ ... a FE space

Consider finite dimensional subspaces

$$\mathcal{S} = \operatorname{span}\{\psi_1, \cdots, \psi_s\} \subset L^2(I)$$
$$\mathcal{Y} = \operatorname{span}\{\nu_1, \cdots, \nu_q\} \subset L^2(\Omega)$$

and the product space

$$\mathcal{S} \cdot \mathcal{Y} \subset L^2(I; L^2(\Omega)).$$

Space-Time Spaces

We represent a function

$$\mathbf{x} = \sum_{j=1}^{s} \sum_{i=1}^{q} \mathbf{x}_{i \cdot j} \nu_i \psi_j \in \mathcal{S} \cdot \mathcal{Y}$$

via its matrix of coefficients

$$\mathbf{X} = \left[\mathbf{x}_{i \cdot j} \right]_{i=1,\dots,q}^{j=1,\dots,s} \in \mathbb{R}^{q,s}$$

and vice versa.

Section 2

Optimal Space Time Product Bases

Space-Time Spaces

Lemma

The space-time L^2-orthogonal projection $x = \Pi_{\mathcal{S}\cdot\mathcal{Y}}\bar{x}$ of a function $\bar{x} \in L^2(I; L^2(\Omega))$ onto $\mathcal{X}$ is given as

$$\mathbf{X} = \mathbf{M}_{\mathcal{Y}}^{-1} \begin{bmatrix} ((x, \nu_1\psi_1))_{\mathcal{S}\cdot\mathcal{Y}} & \dots & ((x, \nu_1\psi_s))_{\mathcal{S}\cdot\mathcal{Y}} \\ \vdots & \ddots & \vdots \\ ((x, \nu_q\psi_1))_{\mathcal{S}\cdot\mathcal{Y}} & \dots & ((x, \nu_q\psi_s))_{\mathcal{S}\cdot\mathcal{Y}} \end{bmatrix} \mathbf{M}_{\mathcal{S}}^{-1},$$

where

$$((x, \nu_i\psi_j))_{\mathcal{S}\cdot\mathcal{Y}} := ((x, \nu_i)_{\mathcal{Y}}, \psi_j)_{\mathcal{S}} := \int_I \Big(\int_\Omega x(\xi, \tau)\nu_i(\xi)\, \mathrm{d}\xi \Big) \psi_j(\tau)\, \mathrm{d}\tau.$$

with the mass matrices

$$\mathbf{M}_{\mathcal{S}} = \big[(\psi_i, \psi_j)_{L^2}\big]_{i,j=1,\dots,s} \quad \text{and} \quad \mathbf{M}_{\mathcal{Y}} = \big[(\nu_i, \nu_j)_{L^2}\big]_{i,j=1,\dots,q}$$

Lemma (Space-time discrete L^2-product)

Let $x^1, x^2 \in \mathcal{S} \cdot \mathcal{Y}$. Then, with

$$\mathbf{x}^\ell = [\mathbf{x}^\ell_{1\cdot 1}, \dots, \mathbf{x}^\ell_{q\cdot 1},\ \mathbf{x}^\ell_{1\cdot 2}, \dots, \mathbf{x}^\ell_{q\cdot 2},\ \dots,\ \mathbf{x}^\ell_{1\cdot s}, \dots, \mathbf{x}^\ell_{q\cdot s}]^\mathsf{T} =: \mathrm{vec}(\mathbf{X}^\ell),$$

the inner product in $\mathcal{S} \cdot \mathcal{Y}$ is given as

$$((x^1, x^2))_{\mathcal{S}\cdot\mathcal{Y}} = \int_I \int_\Omega x^1 x^2 \,\mathrm{d}\xi \,\mathrm{d}\tau = (\mathbf{x}^1)^\mathsf{T} (\mathbf{M}_\mathcal{S} \otimes \mathbf{M}_\mathcal{Y})\, \mathbf{x}^2$$

and the induced norm as

$$\|x^\ell\|^2_{\mathcal{S}\cdot\mathcal{Y}} = \|\mathbf{x}^\ell\|^2_{\mathbf{M}_\mathcal{S}\otimes\mathbf{M}_\mathcal{Y}} = \|\mathbf{M}_\mathcal{Y}^{1/2}\mathbf{X}^\ell\mathbf{M}_\mathcal{S}^{1/2}\|^2_F,$$

$\ell = 1, 2.$

Lemma (Optimal low-rank bases in space)

Given $x \in \mathcal{S} \cdot \mathcal{Y}$ and the associated matrix of coefficients $\mathbf{X}$. The best-approximating subspace $\hat{\mathcal{Y}}$ in the sense that $\|\Pi_{\mathcal{S}\cdot\hat{\mathcal{Y}}} x - x\|_{\mathcal{S}\cdot\mathcal{Y}}$ is minimal over all subspaces of $\mathcal{Y}$ of dimension $\hat{q}$ is given as $\mathrm{span}\{\hat{\nu}_i\}_{i=1,\dots,\hat{q}}$, where

$$\begin{bmatrix} \hat{\nu}_1 \\ \hat{\nu}_2 \\ \vdots \\ \hat{\nu}_{\hat{q}} \end{bmatrix} = V_{\hat{q}}^{\mathsf{T}} \mathbf{M}_{\mathcal{Y}}^{-1/2} \begin{bmatrix} \nu_1 \\ \nu_2 \\ \vdots \\ \nu_q \end{bmatrix},$$

where $V_{\hat{q}}$ is the matrix of the $\hat{q}$ leading left singular vectors of

$$\mathbf{M}_{\mathcal{Y}}^{1/2} \mathbf{X} \mathbf{M}_{\mathcal{S}}^{1/2}.$$

The same arguments apply to the transpose of **X**:

Lemma (Optimal low-rank bases in time[1])

Given $x \in \mathcal{S} \cdot \mathcal{Y}$ *and the associated matrix of coefficients* $\mathbf{X}$*. The best-approximating subspace* $\hat{\mathcal{S}}$ *in the sense that* $\|\Pi_{\hat{\mathcal{S}}\cdot\mathcal{Y}} x - x\|_{\mathcal{S}\cdot\mathcal{Y}}$ *is minimal over all subspaces of* $\mathcal{S}$ *of dimension* $\hat{s}$ *is given as* $\operatorname{span}\{\hat{\psi}_j\}_{j=1,\dots,\hat{s}}$*, where*

$$\begin{bmatrix} \hat{\psi}_1 \\ \hat{\psi}_2 \\ \vdots \\ \hat{\psi}_{\hat{s}} \end{bmatrix} = U_{\hat{s}}^\top \mathbf{M}_{\mathcal{S}}^{-1/2} \begin{bmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_s \end{bmatrix},$$

where $U_{\hat{s}}$ *is the matrix of the* $\hat{s}$ *leading right singular vectors of*

$$\mathbf{M}_{\mathcal{Y}}^{1/2} \mathbf{X} \mathbf{M}_{\mathcal{S}}^{1/2}.$$

[1] see Baumann&PB&JH '16: *ArXiv:1611.04050*

Section 2

Optimal Space Time Product Bases

Relation to POD

The solution of a spatially discretized PDE

$$x\colon \tau \mapsto \mathbb{R}^q$$

is projected to $\mathcal{S} \cdot \mathbb{R}^q$ via

$$\Pi_{\mathcal{S}\cdot\mathcal{Y}} x = \begin{bmatrix} (x_1, \psi_1)_{L^2} & \dots & (x_1, \psi_s)_{L^2} \\ \vdots & \ddots & \vdots \\ (x_q, \psi_1)_{L^2} & \dots & (x_q, \psi_s)_{L^2} \end{bmatrix} \mathbf{M}_{\mathcal{S}}^{-1}.$$

In the (degenerated) case that ψ_j is a delta distribution centered at $\tau_j \in I$, the coefficient matrix degenerates to

$$\begin{bmatrix} x_1(\tau_1) & \dots & x_1(\tau_s) \\ \vdots & \ddots & \vdots \\ x_q(\tau_1) & \dots & x_q(\tau_s) \end{bmatrix}$$

– the standard POD snapshot matrix.

Now consider, in addition, uncertainty, i.e. solutions

$$x(\tau, \xi, w(\omega))$$

with $\tau \in I \subset \mathbb{R}$ ... the time variable
$\xi \in \Omega \subset \mathbb{R}^n$ ... the spatial variable
w ... a random variable

and its numerical approximation:

$$\mathbf{x} \in \mathcal{S} \cdot \mathcal{Y} \cdot \mathcal{W}$$

with $\mathcal{S} \subset L^2(I)$ ... discretized time
$\mathcal{Y} \subset L^2(\Omega)$ ... a FE space
$\mathcal{W}$ ... from *Polynomial Chaos Expansion*

Task: Also, find $\hat{\mathcal{S}} \cdot \hat{\mathcal{Y}} \cdot \hat{\mathcal{W}}$ of much smaller dimension to express $\mathbf{x}$.

Again, we represent a function

$$\mathbf{x} = \sum_{j=1}^{s} \sum_{i=1}^{q} \sum_{\ell=1}^{w} \mathbf{x}_{i\cdot j\cdot \ell}\, \nu_i\, \psi_j\, \eta_\ell \in \mathcal{S} \cdot \mathcal{Y} \cdot \mathcal{W}$$

via its tensor of coefficients

$$\mathbf{X} = \left[\mathbf{x}_{i\cdot j\cdot \ell}\right]_{\substack{i=1,\dots,q\\ j=1,\dots,s\\ \ell=1,\dots,w}} \in \mathbb{R}^{q,s,w}$$

and vice versa.

Optimal Space-Time-Parameter Bases

What is the norm of this tensor product space?[2] – Here, one may base on the expected value

$$\mathbb{E}\|x\|^2_{\mathcal{S}\cdot\mathcal{Y}} = \int_\Sigma \int_I \int_\Omega x(\tau,\xi,\omega)\,\mathrm{d}\xi\,\mathrm{d}\tau\,\mathrm{d}\mathbb{P}(\omega)$$

to get for

$$\mathcal{S}\cdot\mathcal{Y}\cdot\mathcal{W} \ni \mathbf{x} = \sum_{j=1}^{s}\sum_{i=1}^{q}\sum_{\ell=1}^{w} \mathbf{x}_{i\cdot j\cdot \ell}\,\nu_i\,\psi_j\,\eta_\ell$$

that

$$\begin{aligned}\|\mathbf{x}\|^2_{\mathcal{S}\cdot\mathcal{Y}\cdot\mathcal{W}} &= \int_\Sigma \int_I \int_\Omega x(\tau,\xi,\omega)\,\mathrm{d}\xi\,\mathrm{d}\tau\,\mathrm{d}\mathbb{P}(\omega) \\ &= \mathrm{vec}(\mathbf{X})^\mathsf{T}\,(\mathbf{M}_\mathcal{S}\otimes\mathbf{M}_\mathcal{Y}\otimes\mathbf{M}_\mathcal{W})\,\mathrm{vec}(\mathbf{X}).\end{aligned}$$

[2]Here is the question how to extend this to, say, physical parameters...

As for space and time, that $\hat{w}$-dimensional subspace $\hat{\mathcal{W}} \subset \mathcal{W}$ so that $\mathcal{S} \cdot \mathcal{Y} \cdot \hat{\mathcal{W}}$ optimally approximates a given $\mathbf{x} \in$ is defined via an SVD.

- With

$$\|\mathbf{X}\|^2_{\mathcal{S}\cdot\mathcal{Y}\cdot\mathcal{W}} = \|\operatorname{vec}(\mathbf{X})\|^2_{\mathbf{M}_\mathcal{S}\otimes\mathbf{M}_\mathcal{Y}\otimes\mathbf{M}_\mathcal{W}} = \|\mathbf{M}_\mathcal{W}^{1/2}\mathbf{X}^{\mathcal{W}\cdot\mathcal{S}\mathcal{Y}}(\mathbf{M}_\mathcal{Y}\otimes\mathbf{M}_\mathcal{S})^{1/2}\|_F,$$

  with the *matrization* $\mathbf{X}^{\mathcal{W}\cdot\mathcal{S}\mathcal{Y}}$ of the tensor $\mathbf{X}$ along the dimension of $\mathcal{W}$,
- the optimally approximating subspace is defined by the $\hat{w}$ leading left singular vectors of

$$\mathbf{M}_\mathcal{W}^{1/2}\mathbf{X}^{\mathcal{W}\cdot\mathcal{S}\mathcal{Y}}(\mathbf{M}_\mathcal{Y}\otimes\mathbf{M}_\mathcal{S})^{1/2}.$$

[2] see De Lathauwer&De Moor&Vandewalle *A multilinear singular value decomposition*

Section 5

Space-Time Galerkin-POD for Optimal Control

Target: A Space-time Heart Shape

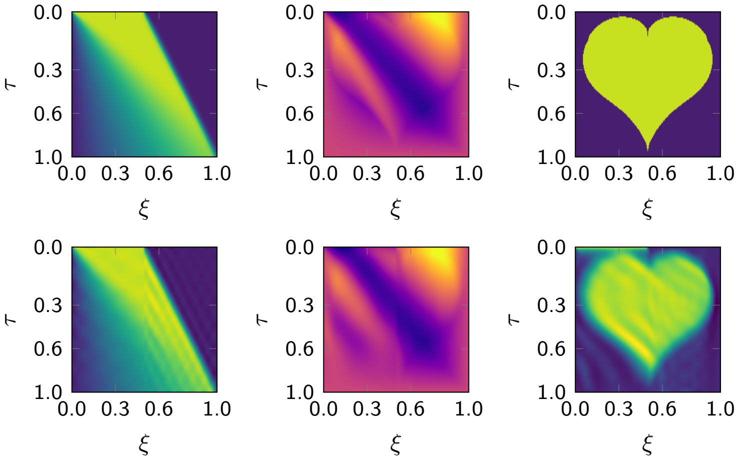


Figure: Illustration of the state, the adjoint, and the target and their approximation via POD-reduced space-time bases.

For a target trajectory $x^* \in L^2(I; L^2(\Omega))$ and a penalization parameter $\alpha > 0$, consider

$$\mathcal{J}(x,u) := \frac{1}{2}\|x - x^*\|_{L^2}^2 + \frac{\alpha}{2}\|u\|_{L^2}^2 \to \min_{u \in L^2(I;L^2(\Omega))}$$

subject to the generic PDE

$$\dot{x} - \Delta x + N(x) = f + u, \quad x(0) = 0. \tag{FWD}$$

If the nonlinearity is smooth, then necessary optimality conditions for (x,u) are given through $u = \frac{1}{\alpha}\lambda$, where λ solves the adjoint equation

$$-\dot{\lambda} - \Delta\lambda + D_x N(x)^{\mathsf{T}}\lambda + x = x^*, \quad \lambda(T) = 0. \tag{BWD}$$

Algorithm (space-time-pod):

Offline Phase

1. Do standard forward/backward solves to compute the matrix of measurements for x and λ.
2. Compute optimal low-dimensional spaces $\hat{\mathcal{S}}$, $\hat{\mathcal{R}}$, $\hat{\mathcal{Y}}$, and $\hat{\Lambda}$ for the space and time discretization of the state x and the adjoint state λ.

Online Phase

3. Solve the space-time Galerkin projected necessary optimality conditions (FWD)-(BWD)[3] for the reduced costate $\hat{\boldsymbol{\lambda}}$.

Evaluation

$\rightarrow$ Inflate $\hat{\mathbf{u}} := \frac{1}{\alpha}\hat{\boldsymbol{\lambda}}$ and apply it in the full order simulation.

[3](FWD)-(BWD) is a two-point boundary value problem with initial and terminal conditions for which time stepping schemes like RKM do not apply.

Numerical Setup

The PDE

- 1D Burger's equation
- $I = (0, 1]$, $\Omega = (0, 1)$
- Viscosity: $\nu = 5 \cdot 10^{-3}$
- Stepfunction as initial value
- Zero Dirichlet conditions

The optimization

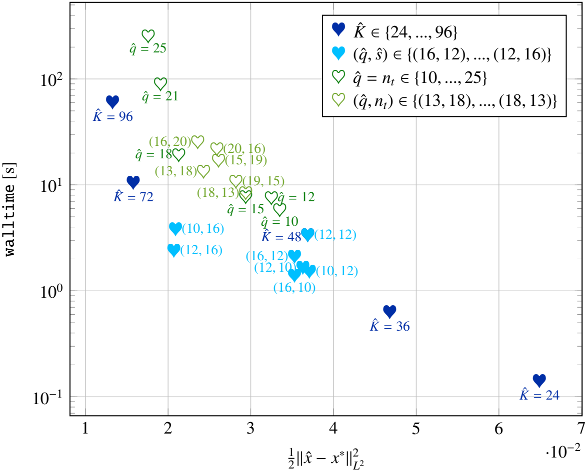
- $\alpha = 10^{-3}$ (space-time-pod)

The full model

- Equidistant space and time grids
- $\mathcal{S} = \mathcal{R}$ ... 120 linear hat functions
- $\mathcal{Y} = \Lambda$ ... 220 linear hat functions

The reduced model

- $\hat{\mathcal{Y}} = \hat{\Lambda}$ ... of dimension $\hat{q} = \hat{p}$
- $\hat{\mathcal{S}} \neq \hat{\mathcal{R}}$ ... of dimensions $\hat{s} = \hat{r}$
- $\hat{q}$, $\hat{p}$, $\hat{s}$, $\hat{r}$ ... varying



Caption:

The achieved tracking vs. the time needed to compute the suboptimal controls by means of

♡, ♡ ... sqp-pod

♥, ♥ ... space-time-pod.

Parameters:

$$\hat{K} :\leftrightarrow \hat{q}, \hat{p}, \hat{r}, \hat{s} = \frac{\hat{K}}{4}$$

$$(\hat{q}, \hat{s}) = (\hat{p}, \hat{r})$$

Conclusion

- The space-time Galerkin POD approach allows for
  - construction of optimized Galerkin bases in space and time
  - in a functional analytical framework
- The resulting space-time Galerkin discretization
  - approximates PDEs by a small system of algebraic equations
  - and naturally extends to boundary value problems in time
  - can be used for efficient computations of (sub)optimal controls

- Future work:
  - Use the functional analytical framework for error estimates.
  - Exploit the freedom of the choice of the measurement functions in $\mathcal{Y}$,
  - to produce, e.g., *optimal* measurements or to compensate for stochastic perturbations.

M. Baumann, P. Benner, and J. Heiland.
Space-Time Galerkin POD with application in optimal control of semi-linear parabolic partial differential equations.
SIAM J. Sci. Comput.(40). 2018.

M. Baumann, J. Heiland, and M. Schmidt.
Discrete input/output maps and their relation to POD.
In P. Benner et al., editors, *Numerical Algebra, Matrix Theory, Differential-Algebraic Equations and Control Theory*. Springer, 2015.

J. Heiland and M. Baumann.
spacetime-galerkin-pod-bfgs-tests – Python/Matlab implementation space-time POD and BFGS for optimal control of Burgers equation.
2016, doi:10.5281/zenodo.166339.

Thank you for your attention!

I am always open for discussion

heiland@mpi-magdeburg.mpg.de
www.janheiland.de
github.com/highlando